

Análisis de Ventas y Gastos en una Empresa con 5 Tiendas



Nombre: Samuel Arteaga López
Módulo: Sistemas de Big Data
Fecha de entrega: 6/02/2025

Índice

1. Descripción.....	3
2. Objetivos del Programa.....	3
1. Cálculo de Totales.....	3
2. Análisis de los Meses Clave.....	3
3. Estadísticas Mensuales.....	4
4. Conclusiones.....	4
3. Requisitos Técnicos.....	4

1. Descripción

Una empresa cuenta con 5 tiendas, y se desea analizar sus ventas y gastos mensuales durante un año (12 meses). Los gastos están divididos en las siguientes categorías:

- Material
- Personal
- Publicidad
- Gastos eléctricos

Deberéis desarrollar un programa en R que permita analizar los datos generados de manera lógica (sin valores extremadamente dispares).

2. Objetivos del Programa

1. Cálculo de Totales

- A nivel global (sumando todas las tiendas):
 - Total de ingresos mensuales.
 - Total de cada tipo de gasto mensual.
 - Beneficio mensual total (beneficio = ingresos - gastos).
- A nivel de cada tienda:
 - Ingresos mensuales de cada tienda.
 - Total de cada tipo de gasto mensual por tienda.
 - Beneficio mensual por tienda.

2. Análisis de los Meses Clave

- A nivel global:
 - Mes con mayores ingresos y mes con menores ingresos.
 - Mes con mayor beneficio y mes con menor beneficio.
- A nivel de cada tienda:
 - Mes con mayores ingresos y mes con menores ingresos por tienda.
 - Mes con mayor beneficio y mes con menor beneficio por tienda.

3. Estadísticas Mensuales

- A nivel global:
 - Media, mediana, moda y desviación típica de los ingresos mensuales.
 - Media, mediana, moda y desviación típica del beneficio mensual.
- A nivel de cada tienda:
 - Media, mediana, moda y desviación típica de los ingresos por tienda.
 - Media, mediana, moda y desviación típica del beneficio por tienda.

4. Conclusiones

- Explicar las conclusiones extraídas de cada análisis.
- Comparar los resultados entre tiendas y determinar qué tiendas tienen mejor rendimiento.
- Identificar patrones o tendencias en las ventas y los gastos tanto a nivel global como por tienda.

3. Requisitos Técnicos

- Los datos deben ser generados dentro del programa con valores lógicos y realistas.
- Se deben usar funciones de R para calcular las métricas estadísticas solicitadas.
- Se debe presentar un resumen de los hallazgos al final del análisis

```
set.seed(123) # Para reproducibilidad de datos
```

```
# Definir nombres de tiendas y meses
```

```
tiendas <- c("Tienda de muebles", "Tienda de airsoft", "Tienda de maquetas",  
"Tienda de ropa", "Tienda gaming")
```

```
meses <- month.name
```

Utilizamos seed() para que genere siempre los mismos valores aleatorios, creamos un array con las 5 tiendas y en la variable meses almacenamos los nombres de los meses del año

```
# Generación de datos lógicos para 5 tiendas en 12 meses
```

```
datos <- data.frame(  
  Mes = rep(meses, each = 5),  
  Tienda = rep(tiendas, times = 12),  
  Ingresos = round(runif(60, 50000, 150000), 2),  
  Material = round(runif(60, 5000, 20000), 2),  
  Personal = round(runif(60, 20000, 50000), 2),  
  Publicidad = round(runif(60, 3000, 10000), 2),  
  GastosElectricos = round(runif(60, 2000, 7000), 2)  
)
```

La función rep() permite repetir, en este caso los meses del año 5 veces, por otro lado se encuentra la función runif(), o mejor, random uniform crea números aleatorios uniformemente

```
# Cálculo de total de gastos y beneficio
```

```
datos$TotalGastos <- datos$Material + datos$Personal + datos$Publicidad +  
datos$GastosElectricos  
datos$Beneficio <- datos$Ingresos - datos$TotalGastos
```

Creamos dos nuevas columnas dentro del dataframe datos que almacena la suma del total de gastos y el beneficio

```
# Cálculo de totales globales
```

```
totalesGlobales <- aggregate(cbind(Ingresos, Total_Gastos, Beneficio, Material,  
Personal, Publicidad, GastosElectricos) ~ Mes, data = datos, sum)
```

A continuación, se agrupan los ingresos, gastos y beneficio por mes y se suman

```
# Cálculo de totales por tienda
```

```
totalesTienda <- aggregate(cbind(Ingresos, TotalGastos, Beneficio, Material,  
Personal, Publicidad, GastosElectricos) ~ Tienda + Mes, data = datos, sum)
```

Igual que en la casuística anterior, se agrupan los ingresos, gastos y beneficio pero se agrupan por tienda y mes y se suman

```
# Identificación de meses clave
ingresosMaximosMes <- totalGlobales$Mes[which.max(totalGlobales$Ingresos)]
ingresosMinimosMes <- totalGlobales$Mes[which.min(totalGlobales$Ingresos)]
beneficioMaximoMes <- totalGlobales$Mes[which.max(totalGlobales$Beneficio)]
beneficioMinimoMes <- totalGlobales$Mes[which.min(totalGlobales$Beneficio)]
```

Para sacar los ingresos y beneficios maximos y minimos obtendremos, de todo el total por cada mes, los ingresos y beneficios mas altos y bajos con las funciones min() max() dentro del filtrador which

```
# Meses clave por tienda
mesesClaveTienda <- aggregate(cbind(Ingresos, Beneficio) ~ Tienda + Mes, data =
datos, sum)
ingresosMaxTienda <- mesesClaveTienda[which.max(mesesClaveTienda$Ingresos),]
ingresosMinTienda <- mesesClaveTienda[which.min(mesesClaveTienda$Ingresos),]
beneficioMaxTienda <- mesesClaveTienda[which.max(mesesClaveTienda$Beneficio),]
beneficioMinTienda <- mesesClaveTienda[which.min(mesesClaveTienda$Beneficio),]
```

Igual que hemos hecho para todo el total lo hacemos para cada tienda en especifico obteniendo el maximo y minimo de cada tienda agrupado por esta tienda y el mes

```
# Función para calcular la moda
calcularModa <- function(x) {
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]
}
```

Creamos una funcion de calcular moda donde primero nos quedamos con los valores que no estén repetidos, a continuación, con tabulate obtenemos el valor que más se repite el cual se obtiene mediante match, donde x y ux son los mismos elementos dentro del dataframe

```
# Estadísticas globales
ingresosEstadisticas <- c(
  mean = mean(totalGlobales$Ingresos),
  median = median(totalGlobales$Ingresos),
  moda = calcularModa(totalGlobales$Ingresos),
  sd = sd(totalGlobales$Ingresos)
)
beneficioEstadisticas <- c(
  mean = mean(totalGlobales$Beneficio),
  median = median(totalGlobales$Beneficio),
  moda = calcularModa(totalGlobales$Beneficio),
  sd = sd(totalGlobales$Beneficio)
)
```

Seguidamente dentro del array ingresos y beneficios estadísticos creamos la media, la mediana, la moda con la función anteriormente comentada y la desviación estándar de los totales globales

```
# Estadísticas por tienda
estadisticasTienda <- aggregate(cbind(Ingresos, Beneficio) ~ Tienda, data = datos,
function(x) c(mean = mean(x), median = median(x), moda = calcularModa(x), sd =
sd(x)))
```

Lo mismo para las tiendas pero en este caso creamos una función que nos devuelve en un array la media, y el resto de valores agrupados de cada tienda

A continuación imprimimos todos los resultados del análisis de las cinco tiendas

```
# Resumen de hallazgos
cat("Mes con mayores ingresos:", ingresosMaximosMes, "\n")
cat("Mes con menores ingresos:", ingresosMinimosMes, "\n")
cat("Mes con mayor beneficio:", beneficioMaximoMes, "\n")
cat("Mes con menor beneficio:", beneficioMinimoMes, "\n")
```

```
Mes con mayores ingresos: May
Mes con menores ingresos: September
Mes con mayor beneficio: May
Mes con menor beneficio: September
```

```
cat("\nEstadísticas de ingresos mensuales:\n")
print(ingresosEstadisticas)
```

```
cat("\nEstadísticas de beneficio mensual:\n")
print(beneficioEstadisticas)
```

```
Estadísticas de ingresos mensuales:
  mean   median    moda    sd
502760.26 491363.79 497039.65 69208.41
```

```
Estadísticas de beneficio mensual:
  mean   median    moda    sd
212129.53 188587.89 181997.84 77986.31
```

```
cat("\nEstadísticas por tienda:\n")
print(estadisticasTienda)
```

```
Estadísticas por tienda:
      Tienda Ingresos.mean Ingresos.median Ingresos.moda Ingresos.sd
1  Tienda de airsoft    101089.77      99071.99    128830.51    24074.53
2  Tienda de maquetas    104955.90     111732.44     90897.69    24891.88
3   Tienda de muebles     98022.09     88268.68     78757.75    37852.11
4    Tienda de ropa    103198.73     96014.02    138301.74    29419.59
5   Tienda gaming     95493.77     91553.88    144046.73    33721.70
  Beneficio.mean Beneficio.median Beneficio.moda Beneficio.sd
1    43177.54      49456.56      82314.55    26999.88
2    48932.94      52916.25     40090.03    26867.81
3    36779.19      28898.81     10731.72    41661.26
4    44937.21      37347.21     92572.56    32887.95
5    38302.65      26362.59     85022.71    33186.53
```

```
cat("\nMeses clave por tienda:\n")
print(ingresosMaxTienda)
print(ingresosMinTienda)
print(beneficioMaxTienda)
print(beneficioMinTienda)
```

```
Meses clave por tienda:
      Tienda Mes Ingresos Beneficio
44 Tienda de ropa May  149427  93732.51
      Tienda Mes Ingresos Beneficio
30 Tienda gaming July  52461.37  10137.56
      Tienda Mes Ingresos Beneficio
44 Tienda de ropa May  149427  93732.51
      Tienda Mes Ingresos Beneficio
18 Tienda de muebles February 54555.65 -18690.48
```

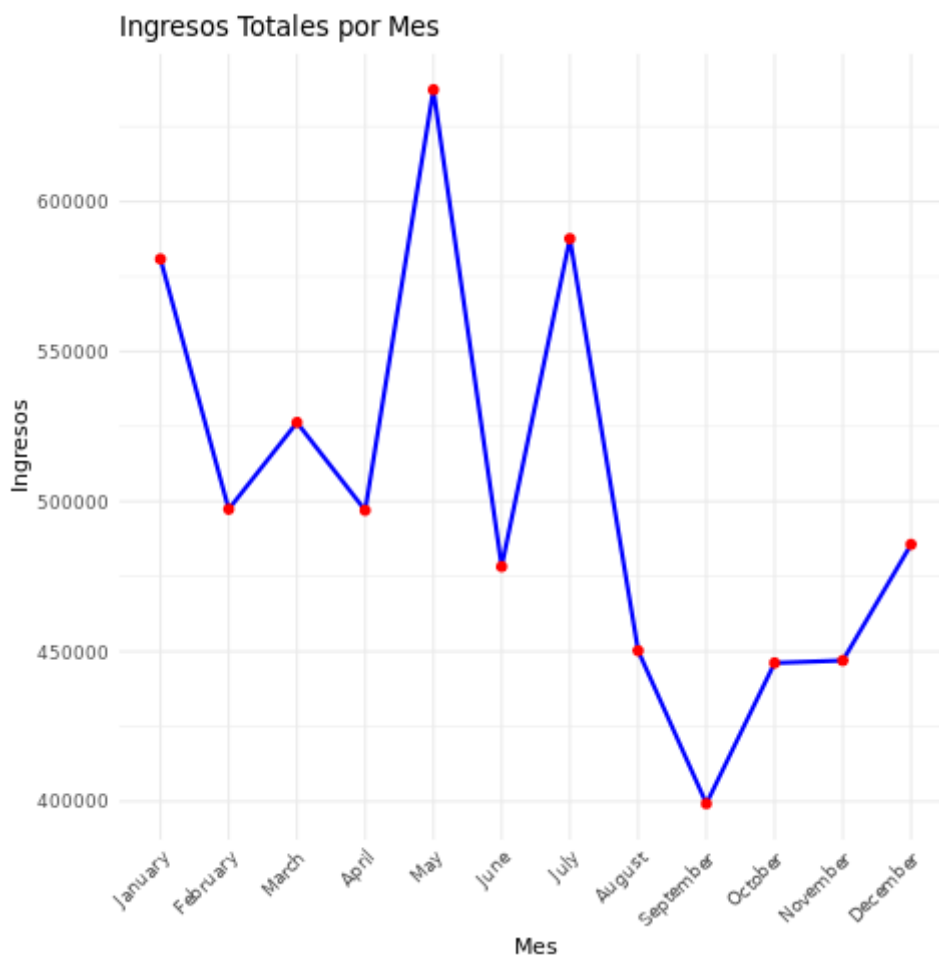
Conclusiones:

Se observa que algunas tiendas tienen ingresos y beneficios más estables, mientras que otras presentan más variabilidad.

Las tendencias en los gastos también afectan directamente el beneficio, especialmente en personal y publicidad.

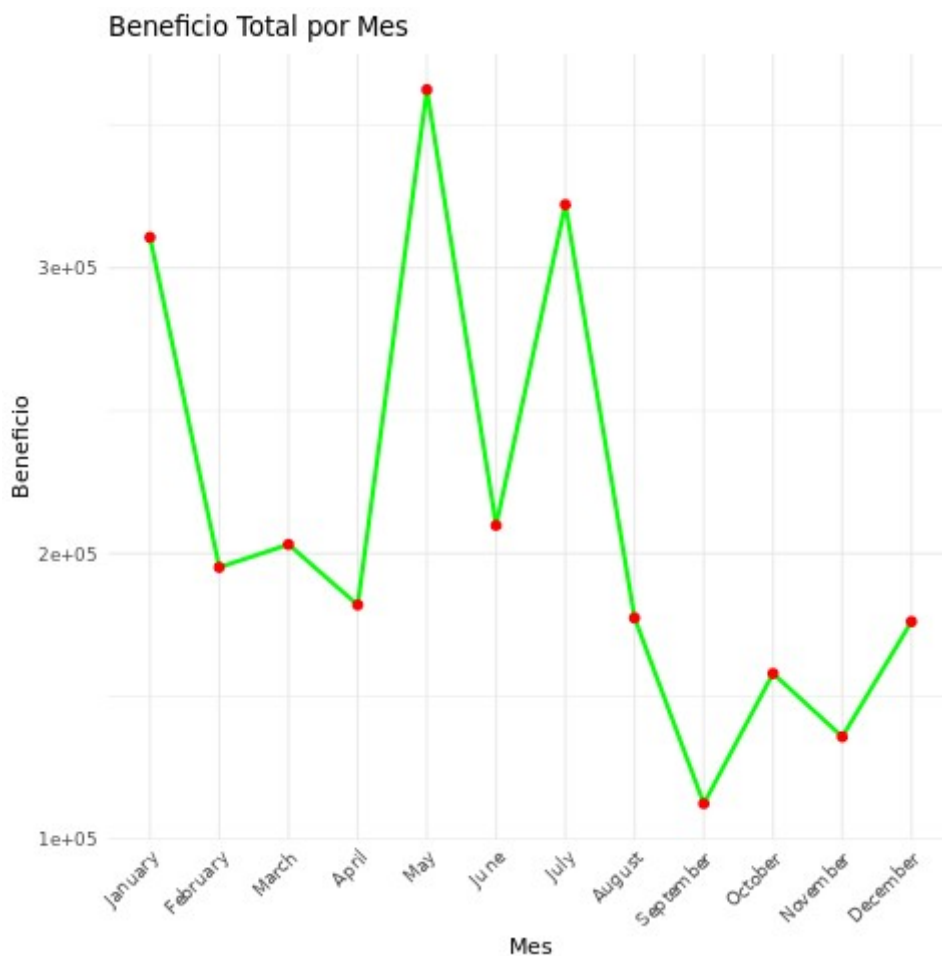
Es recomendable analizar estrategias para optimizar los gastos en las tiendas con menor rentabilidad.


```
# Gráfico de ingresos por mes
ggplot(totalGlobales, aes(x = Mes, y = Ingresos, group = 1)) +
  geom_line(color = "blue", size = 1) +
  geom_point(color = "red", size = 2) +
  ggtitle("Ingresos Totales por Mes") +
  xlab("Mes") +
  ylab("Ingresos") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotar etiquetas
```



Se observa en el grafico la coincidencia de la salida por pantalla donde se comentaba que el mejor mes era Mayo y el peor, Septiembre a nivel global, vemos que los tres mejores junto a Mayo son Enero y Julio, mientras que los peores son, aparte de Septiembre, Agosto, Octubre y Noviembre.

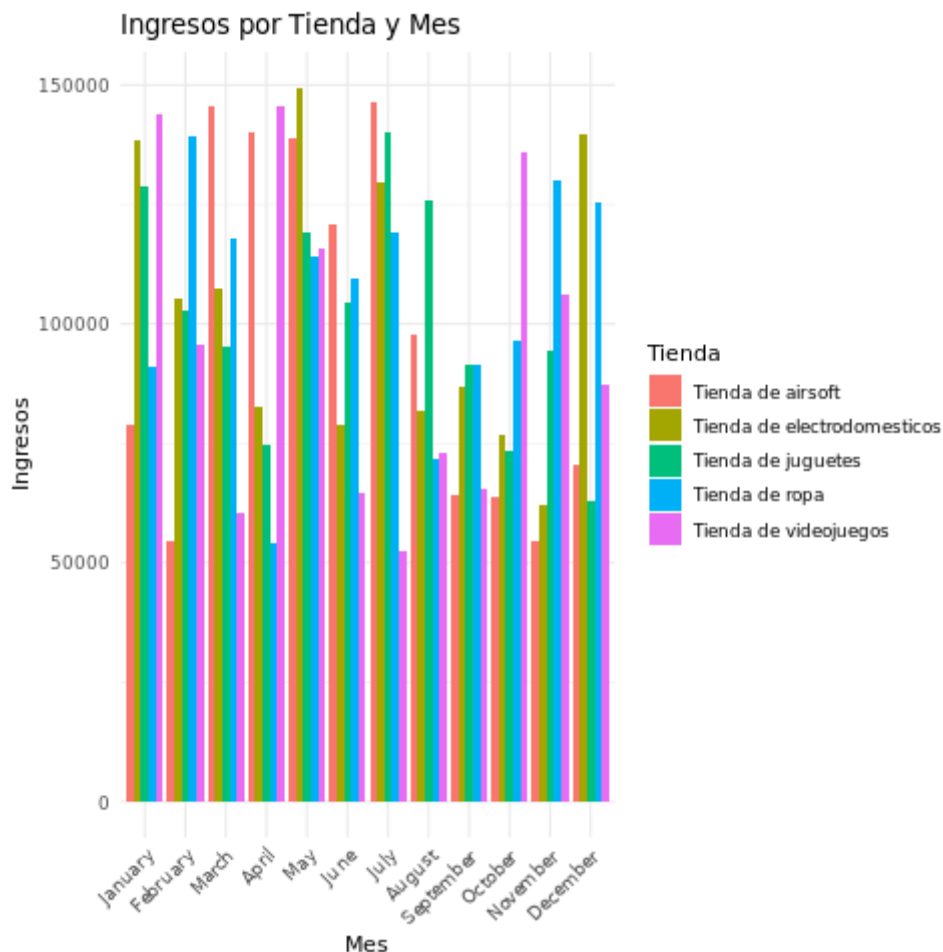
```
# Gráfico de beneficios por mes
ggplot(totalGlobales, aes(x = Mes, y = Beneficio, group = 1)) +
  geom_line(color = "green", size = 1) +
  geom_point(color = "red", size = 2) +
  ggtitle("Beneficio Total por Mes") +
  xlab("Mes") +
  ylab("Beneficio") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotar etiquetas
```



Similar a los ingresos encontramos que los mejores meses de ingresos son Enero, Mayo y Julio, mientras que los peores son Septiembre, Octubre y Noviembre

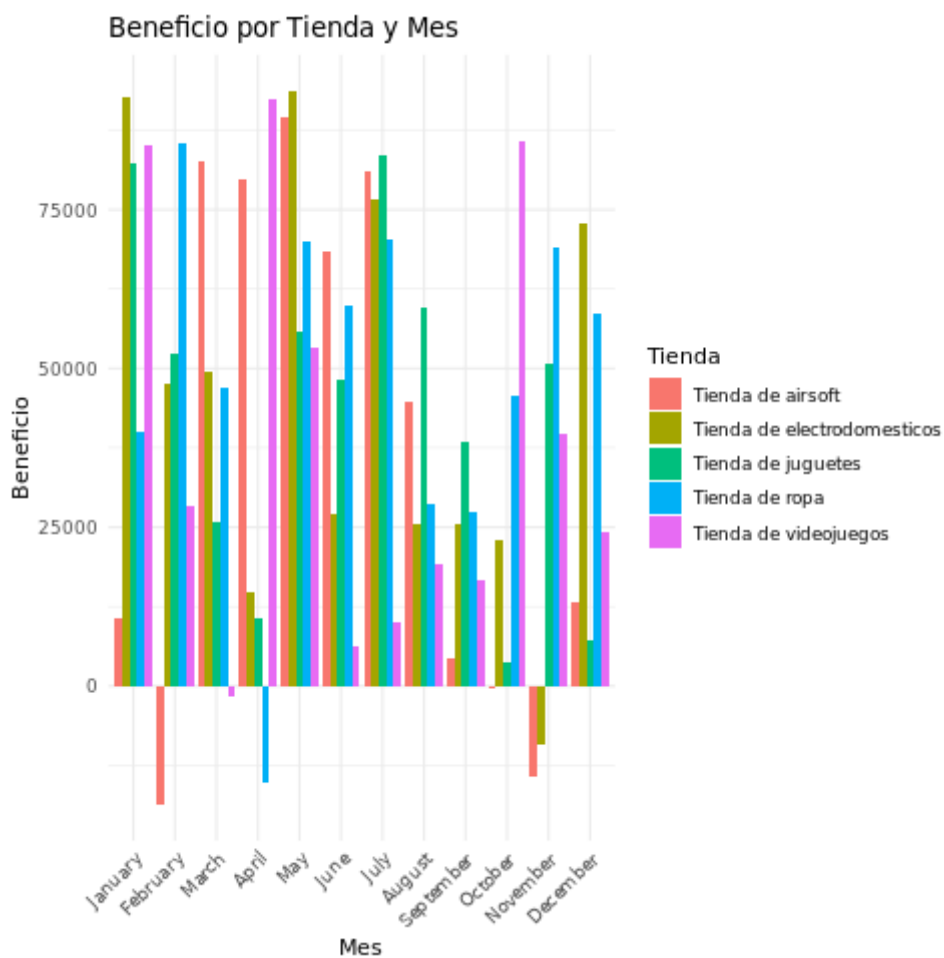
```
# Gráfico de ingresos por tienda
ggplot(totalesTienda, aes(x = Mes, y = Ingresos, fill = Tienda)) +
```

```
geom_bar(stat = "identity", position = "dodge") +  
ggtitle("Ingresos por Tienda y Mes") +  
xlab("Mes") +  
ylab("Ingresos") +  
theme_minimal() +  
theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotar etiquetas
```



En esta gráfica observamos los ingresos por tienda en los meses del año donde cada tienda tiene su valor más alto según su mes, por ejemplo la tienda de airsoft tiene las peores ventas en noviembre pero las mejores en Junio, esto podría deberse a la temporada alta en verano que en invierno por el buen tiempo para jugar.

```
# Gráfico de beneficios por tienda
ggplot(totalesTienda, aes(x = Mes, y = Beneficio, fill = Tienda)) +
  geom_bar(stat = "identity", position = "dodge") +
  ggtitle("Beneficio por Tienda y Mes") +
  xlab("Mes") +
  ylab("Beneficio") +
  theme_minimal() +
  theme(axis.text.x = elementText(angle = 45, hjust = 1)) # Rotar etiquetas
```



Sin embargo en el beneficio, a contraste con los ingresos vemos tiendas que pierden dinero en algunos meses como es el caso del airsoft donde en temporada baja como meses de Febrero y Noviembre tienen beneficio negativo